

# Selected configuration interaction using a transcorrelated Hamiltonian

---

A. Ammar<sup>1</sup>, Y. Damour<sup>1</sup>, E. Giner<sup>2</sup>, P-F. Loos<sup>1</sup>, A. Scemama<sup>1</sup>

08/10/2024

<sup>1</sup>Laboratoire de Chimie et Physique Quantiques, Univ. Toulouse/CNRS

<sup>2</sup>Laboratoire de Chimie Théorique, Univ. Paris Sorbonne/CNRS

## Objective

The most accurate possible approximation to the solution of the Schrödinger equation for (small) molecules, with a reasonable energy footprint.

1. Reaching the full configuration interaction (FCI) accuracy
2. Going beyond FCI

# Full configuration interaction (FCI)

## Exact solution of $\hat{H}\Psi = E\Psi$ in a given N-electron basis

- All possible ways to put  $N_{\uparrow}$  electrons in  $M$  orbitals and  $N_{\downarrow}$  electrons in  $M$  orbitals:

$$\begin{aligned} \Psi = & c_1 \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix} + c_2 \begin{pmatrix} -- \\ \uparrow- \\ -\downarrow \end{pmatrix} + c_3 \begin{pmatrix} \uparrow- \\ -- \\ -\downarrow \end{pmatrix} + c_4 \begin{pmatrix} -- \\ -\downarrow \\ \uparrow- \end{pmatrix} + c_5 \begin{pmatrix} -- \\ \uparrow\downarrow \\ -- \end{pmatrix} + \\ & c_6 \begin{pmatrix} \uparrow- \\ -\downarrow \\ -- \end{pmatrix} + c_7 \begin{pmatrix} -\downarrow \\ -- \\ \uparrow- \end{pmatrix} + c_8 \begin{pmatrix} -\downarrow \\ \uparrow- \\ -- \end{pmatrix} + c_9 \begin{pmatrix} \uparrow\downarrow \\ -- \\ -- \end{pmatrix} \end{aligned}$$

# Full configuration interaction (FCI)

## Exact solution of $\hat{H}\Psi = E\Psi$ in a given N-electron basis

- All possible ways to put  $N_{\uparrow}$  electrons in  $M$  orbitals and  $N_{\downarrow}$  electrons in  $M$  orbitals:

$$\begin{aligned}\Psi = & c_1 \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix} + c_2 \begin{pmatrix} -- \\ \uparrow- \\ -\downarrow \end{pmatrix} + c_3 \begin{pmatrix} \uparrow- \\ -- \\ -\downarrow \end{pmatrix} + c_4 \begin{pmatrix} -- \\ -\downarrow \\ \uparrow- \end{pmatrix} + c_5 \begin{pmatrix} -- \\ \uparrow\downarrow \\ -- \end{pmatrix} + \\ & c_6 \begin{pmatrix} \uparrow- \\ -\downarrow \\ -- \end{pmatrix} + c_7 \begin{pmatrix} -\downarrow \\ -- \\ \uparrow- \end{pmatrix} + c_8 \begin{pmatrix} -\downarrow \\ \uparrow- \\ -- \end{pmatrix} + c_9 \begin{pmatrix} \uparrow\downarrow \\ -- \\ -- \end{pmatrix}\end{aligned}$$

- Variational optimization of  $c_i$

# Full configuration interaction (FCI)

## Exact solution of $\hat{H}\Psi = E\Psi$ in a given N-electron basis

- All possible ways to put  $N_{\uparrow}$  electrons in  $M$  orbitals and  $N_{\downarrow}$  electrons in  $M$  orbitals:

$$\begin{aligned} \Psi = & c_1 \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix} + c_2 \begin{pmatrix} -- \\ \uparrow- \\ -\downarrow \end{pmatrix} + c_3 \begin{pmatrix} \uparrow- \\ -- \\ -\downarrow \end{pmatrix} + c_4 \begin{pmatrix} -- \\ -\downarrow \\ \uparrow- \end{pmatrix} + c_5 \begin{pmatrix} -- \\ \uparrow\downarrow \\ -- \end{pmatrix} + \\ & c_6 \begin{pmatrix} \uparrow- \\ -\downarrow \\ -- \end{pmatrix} + c_7 \begin{pmatrix} -\downarrow \\ -- \\ \uparrow- \end{pmatrix} + c_8 \begin{pmatrix} -\downarrow \\ \uparrow- \\ -- \end{pmatrix} + c_9 \begin{pmatrix} \uparrow\downarrow \\ -- \\ -- \end{pmatrix} \end{aligned}$$

- Variational optimization of  $c_i$
- Intractable :  $\mathcal{O}(N!)$  scaling of parameters (*Quantum computing?*)

$$N_{\text{det}} = \left( \frac{M!}{N_{\uparrow}!(M - N_{\uparrow})!} \right) \times \left( \frac{M!}{N_{\downarrow}!(M - N_{\downarrow})!} \right)$$

# Full configuration interaction (FCI)

## Full Configuration Interaction (FCI)

- Exact solution of  $\hat{H}\Psi = E\Psi$  in a given **N-electron basis** of Slater determinants
- The determinant basis is derived from the **1-electron** basis set
- **Only approximation of FCI** : 1-electron basis-set incompleteness

# Full configuration interaction (FCI)

## Full Configuration Interaction (FCI)

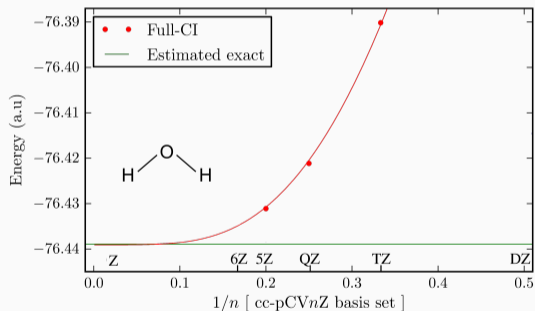
- Exact solution of  $\hat{H}\Psi = E\Psi$  in a given **N-electron basis** of Slater determinants
- The determinant basis is derived from the **1-electron** basis set
- **Only approximation of FCI** : 1-electron basis-set incompleteness

## All post-HF methods are approximations to FCI within the same basis set

- FCI is the reference for benchmarking post-HF methods (not experiment!)
  - Configuration Interaction (CI): Use *not all* possible determinants: variational optimization of  $c_i$
  - Coupled Cluster (CC): Use *all* possible determinants: non-variational optimization of a reduced number of parameters  $t_i$
  - Perturbative methods: Direct approximation of  $c_i$  in one shot

# How accurate is FCI?

## Water molecule



Basis	E(FCI)	Err ( $mE_h$ )	Corr energy
cc-pCVDZ	-76.28214	156.8	58.1 %
cc-pCVTZ	-76.38829	50.7	86.5 %
cc-pCVQZ	-76.41932	19.6	94.8 %
cc-pCV5Z	-76.42855	10.4	97.2 %
Est. exact	-76.43894	0.	100.0 %

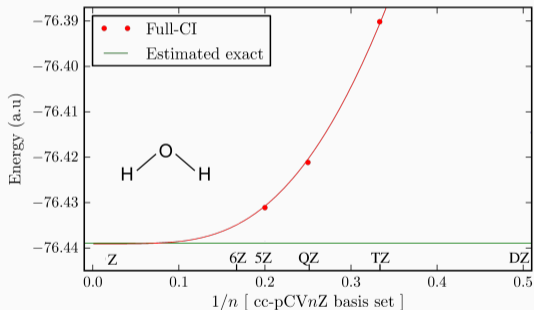
Basis set incompleteness error

Caffarel *et al*, doi:10.1063/1.4947093



# How accurate is FCI?

## Water molecule



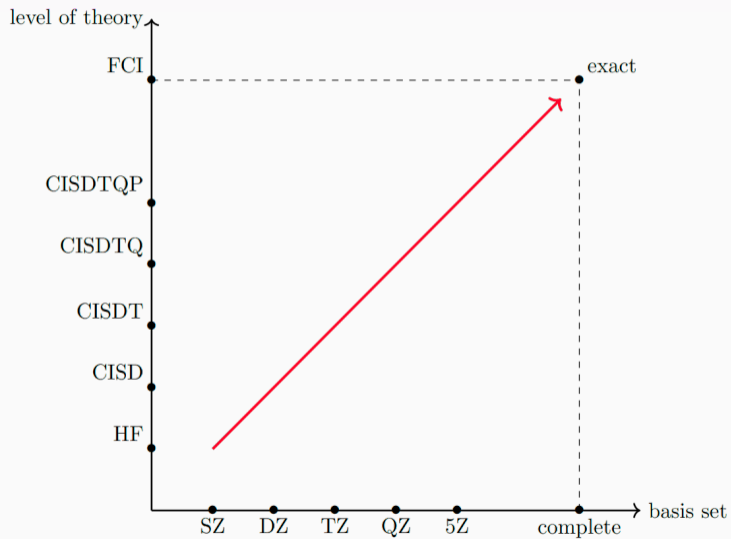
Caffarel *et al*, doi:10.1063/1.4947093

Basis	E(FCI)	Err ( $mE_h$ )	Corr energy
cc-pCVDZ	-76.28214	156.8	58.1 %
cc-pCVTZ	-76.38829	50.7	86.5 %
cc-pCVQZ	-76.41932	19.6	94.8 %
cc-pCV5Z	-76.42855	10.4	97.2 %
Est. exact	-76.43894	0.	100.0 %

### Basis set incompleteness error

Good news: Error cancellation is large in energy *differences*

# Pople diagram



# Obtaining the FCI energy within a given basis set

## Extrapolation

- Extrapolating: CISD, CISDT, CISDTQ, ...
- Extrapolating: CCSD, CCSDT, CCSDTQ, ...

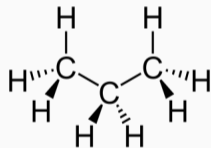
# Obtaining the FCI energy within a given basis set

## Extrapolation

- Extrapolating: CISD, CISDT, CISDTQ, ...
- Extrapolating: CCSD, CCSDT, CCSDTQ, ...

## Brute force

- Propane/STO-3G: 26 electrons in 23 MOs  
(Gao *et al* doi:10.1021/acs.jctc.3c01190)
  - $1.3 \cdot 10^{12}$  Slater determinants
  - 113 hours on 256 nodes (10 240 cores)



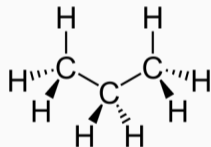
# Obtaining the FCI energy within a given basis set

## Extrapolation

- Extrapolating: CISD, CISDT, CISDTQ, ...
- Extrapolating: CCSD, CCSDT, CCSDTQ, ...

## Brute force

- Propane/STO-3G: 26 electrons in 23 MOs  
(Gao *et al* doi:10.1021/acs.jctc.3c01190)
  - $1.3 \cdot 10^{12}$  Slater determinants
  - 113 hours on 256 nodes (10 240 cores)



## Sparse approaches

1. FCI-QMC: Stochastic solution of FCI equations  
(Booth *et al*, doi:10.1063/1.3193710)
2. Selected CI: CIPSI or SHCI

# CIPSI: CI using a Perturbative Selection made Iteratively

## Algorithm

1. Start with a single determinant in the set  $\mathcal{D}_0$ :  $\Psi_n = \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix}$ .

# CIPSI: CI using a Perturbative Selection made Iteratively

## Algorithm

1. Start with a single determinant in the set  $\mathcal{D}_0$ :  $\Psi_n = \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix}$ .
2. Among all other determinants, estimate with perturbation theory (PT2) which one will be the most important, and add it to the space  $\mathcal{D}_{n+1}$ .

# CIPSI: CI using a Perturbative Selection made Iteratively

## Algorithm

1. Start with a single determinant in the set  $\mathcal{D}_0$ :  $\Psi_n = \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix}$ .
2. Among all other determinants, estimate with perturbation theory (PT2) which one will be the most important, and add it to the space  $\mathcal{D}_{n+1}$ .

$$\forall |\alpha\rangle \in \{\hat{T}_{\text{SD}}|\Psi_n\rangle\} \setminus \{\mathcal{D}_n\}, \text{ compute } e_\alpha = \frac{\langle \alpha | \mathcal{H} | \Psi_n \rangle^2}{E(\Psi_n) - \langle \alpha | \mathcal{H} | \alpha \rangle},$$

and select  $|\alpha\rangle$  with the largest  $|e_\alpha|$ :  $\mathcal{D}_{n+1} = \mathcal{D} \cup \{|\alpha\rangle\}$ .



# CIPSI: CI using a Perturbative Selection made Iteratively

## Algorithm

1. Start with a single determinant in the set  $\mathcal{D}_0$ :  $\Psi_n = \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix}$ .
2. Among all other determinants, estimate with perturbation theory (PT2) which one will be the most important, and add it to the space  $\mathcal{D}_{n+1}$ .
3. Estimate with PT2 the energy contribution coming from all the other determinants:

$$E_{\text{PT2}}(\Psi_n) = \sum_{|\alpha\rangle \notin \mathcal{D}_n} e_\alpha, \quad \text{and} \quad E_{\text{FCI}} = E(\Psi_n) + (1 + \epsilon_n)E_{\text{PT2}}(\Psi_n)$$

# CIPSI: CI using a Perturbative Selection made Iteratively

## Algorithm

1. Start with a single determinant in the set  $\mathcal{D}_0$ :  $\Psi_n = \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix}$ .
2. Among all other determinants, estimate with perturbation theory (PT2) which one will be the most important, and add it to the space  $\mathcal{D}_{n+1}$ .
3. Estimate with PT2 the energy contribution coming from all the other determinants:
4. Minimize the energy  $E(\Psi_{n+1})$  over the space  $\mathcal{D}_{n+1}$ , and obtain a new wave function and energy:  $\Psi_{n+1} = c_1 \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix} + c_2 \begin{pmatrix} -- \\ \uparrow\downarrow \\ -- \end{pmatrix}$ .
5. Go back to step 2.

# CIPSI: CI using a Perturbative Selection made Iteratively

## Algorithm

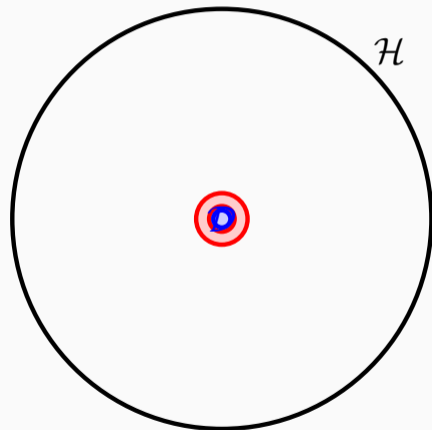
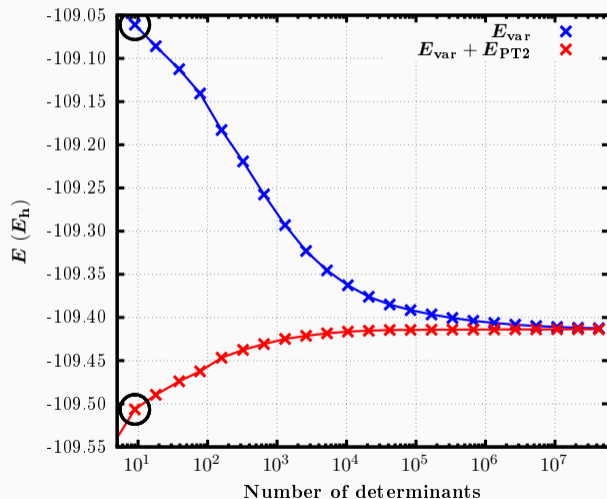
1. Start with a single determinant in the set  $\mathcal{D}_0$ :  $\Psi_n = \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix}$ .
2. Among all other determinants, estimate with perturbation theory (PT2) which one will be the most important, and add it to the space  $\mathcal{D}_{n+1}$ .
3. Estimate with PT2 the energy contribution coming from all the other determinants:
4. Minimize the energy  $E(\Psi_{n+1})$  over the space  $\mathcal{D}_{n+1}$ , and obtain a new wave function and energy:  $\Psi_{n+1} = c_1 \begin{pmatrix} -- \\ -- \\ \uparrow\downarrow \end{pmatrix} + c_2 \begin{pmatrix} -- \\ \uparrow\downarrow \\ -- \end{pmatrix}$ .
5. Go back to step 2.

## Output

Wavefunctions  $\Psi_n$     Variational energies  $E(\Psi_n)$     PT2 corrections  $E_{\text{PT2}}(\Psi_n)$

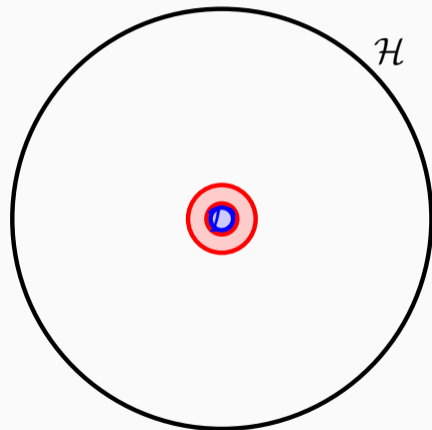
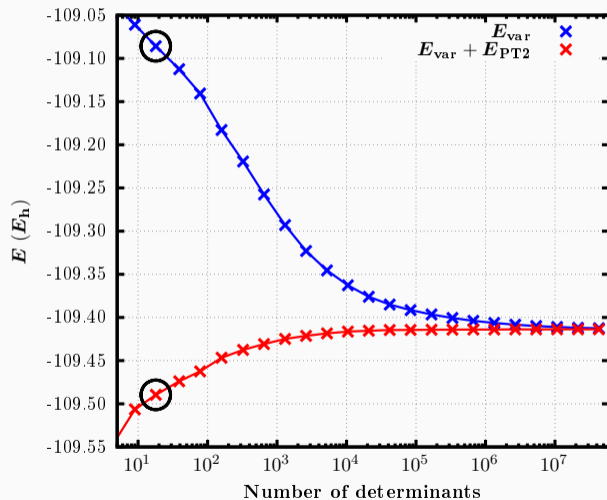
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



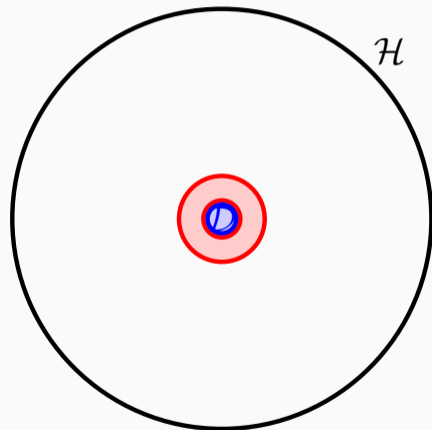
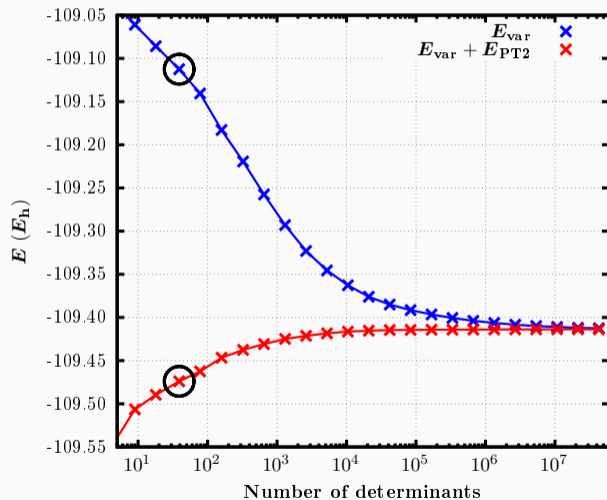
# CIPSI: CI using a Perturbative Selection made Iteratively

## $N_2$ , aug-cc-pVTZ+3s3p3d



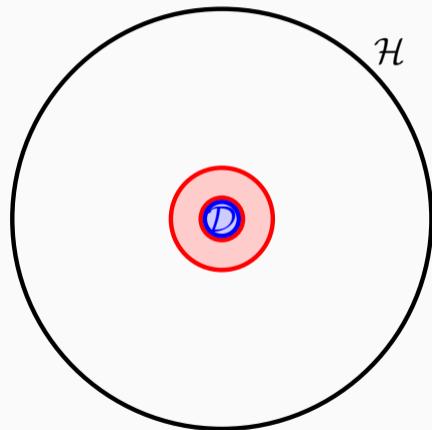
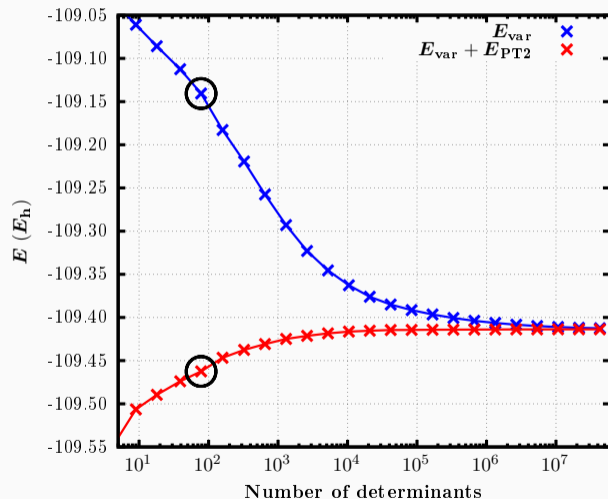
# CIPSI: CI using a Perturbative Selection made Iteratively

## $N_2$ , aug-cc-pVTZ+3s3p3d



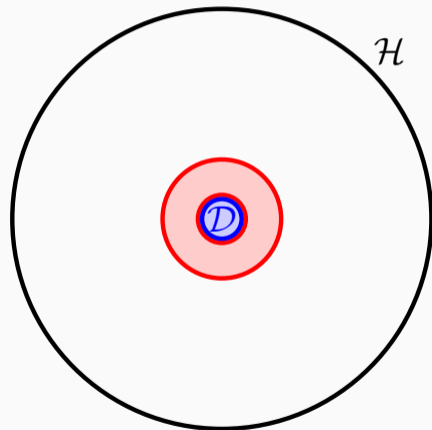
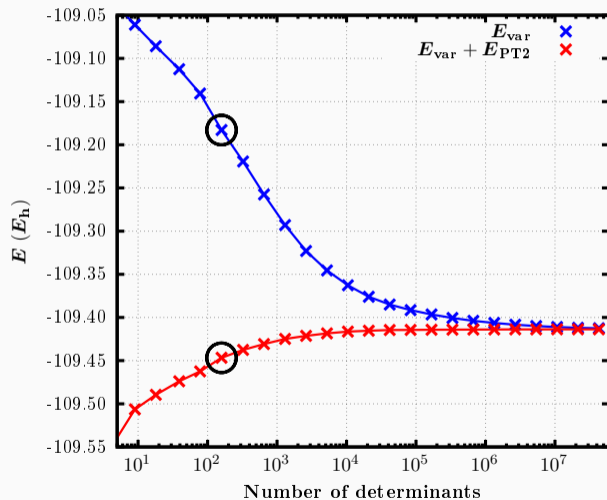
# CIPSI: CI using a Perturbative Selection made Iteratively

## $N_2$ , aug-cc-pVTZ+3s3p3d



# CIPSI: CI using a Perturbative Selection made Iteratively

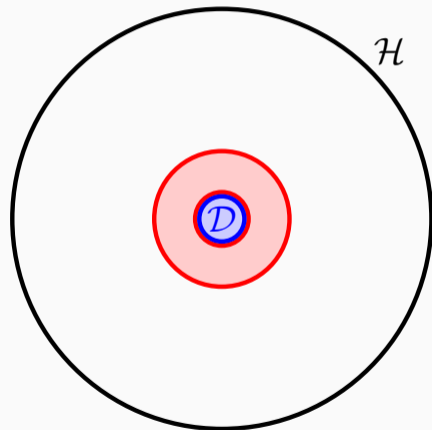
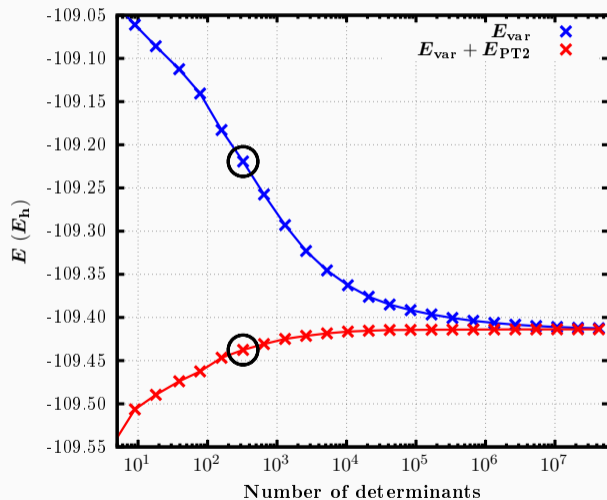
$N_2$ , aug-cc-pVTZ+3s3p3d





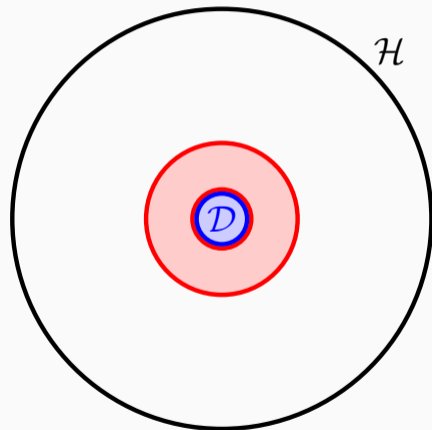
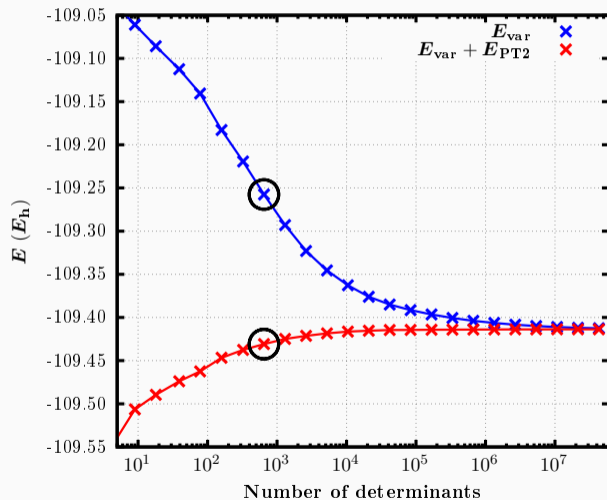
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



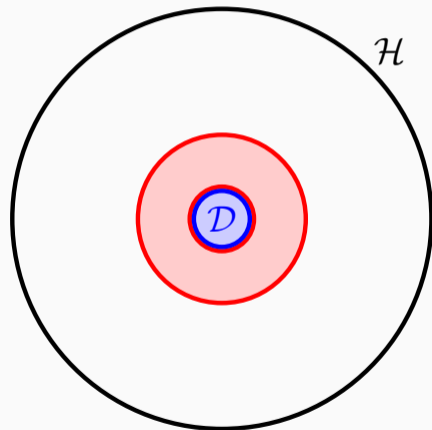
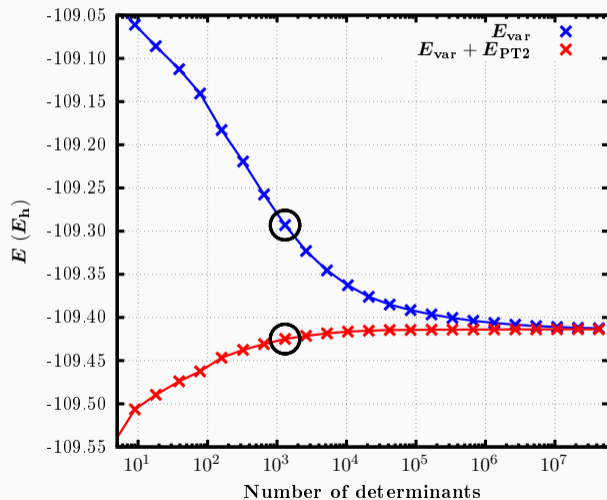
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



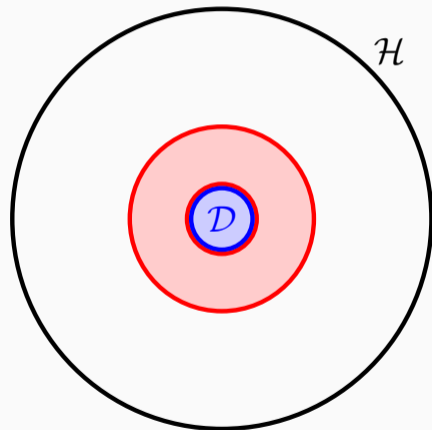
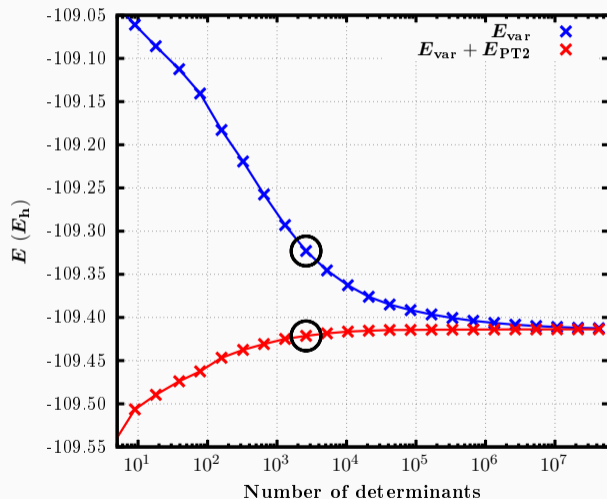
# CIPSI: CI using a Perturbative Selection made Iteratively

## $N_2$ , aug-cc-pVTZ+3s3p3d



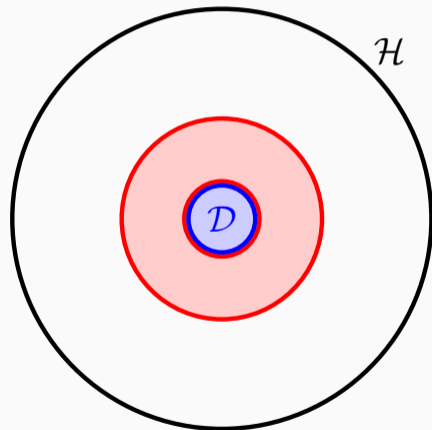
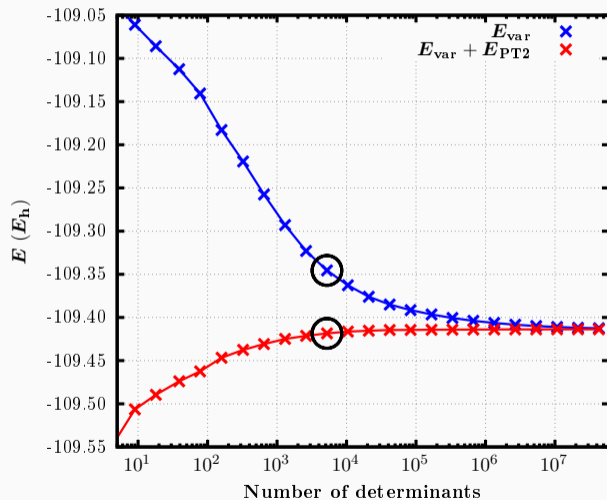
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



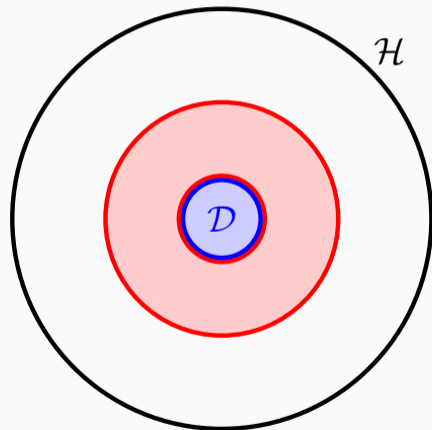
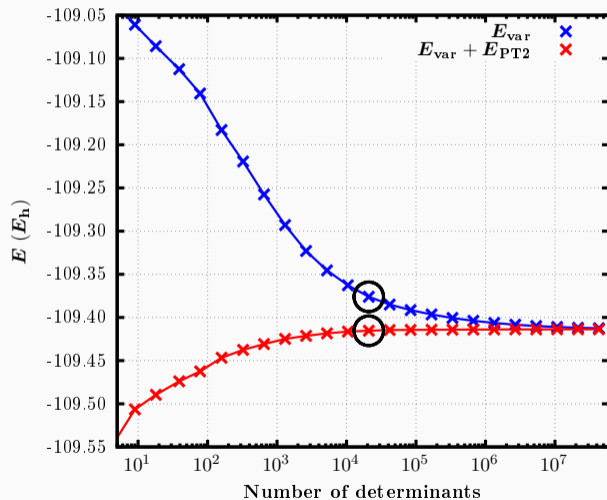
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



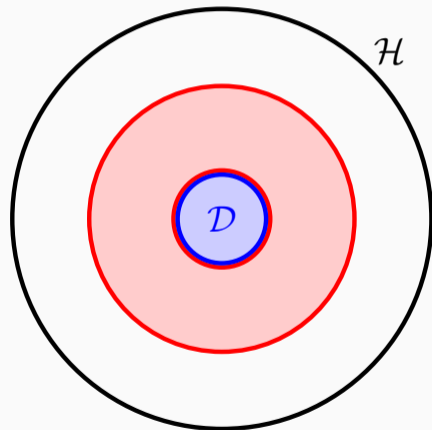
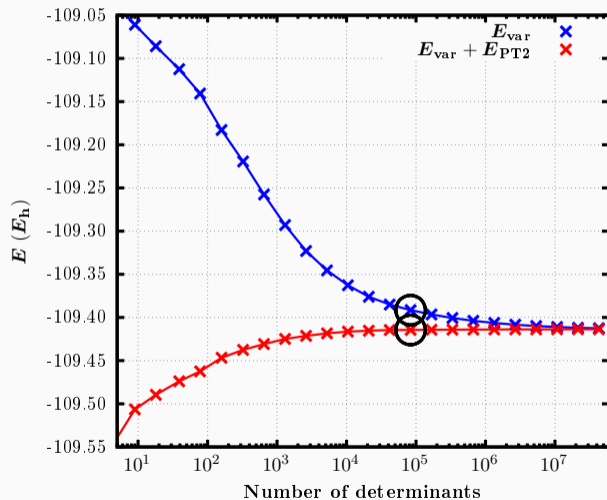
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



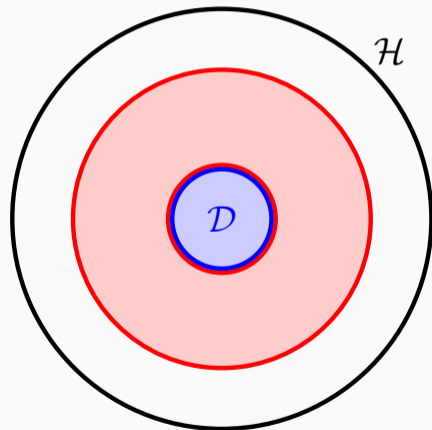
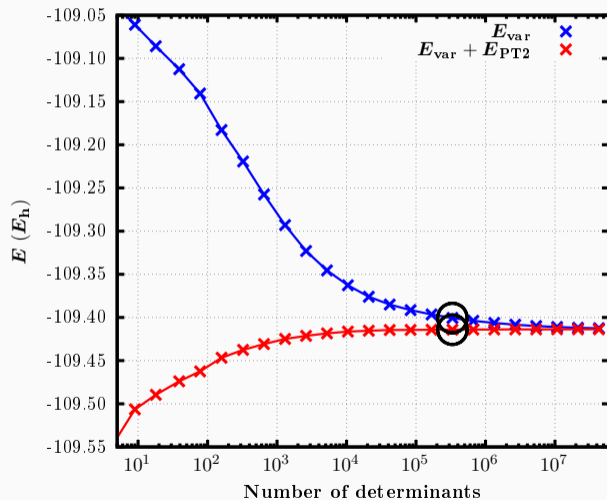
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



# CIPSI: CI using a Perturbative Selection made Iteratively

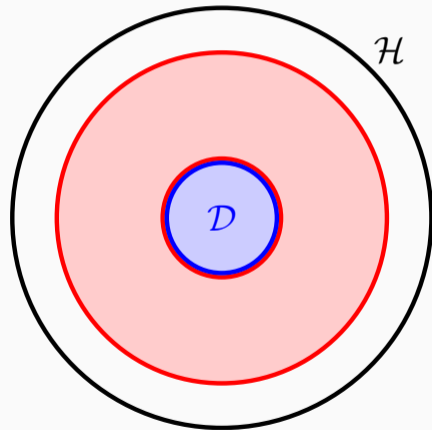
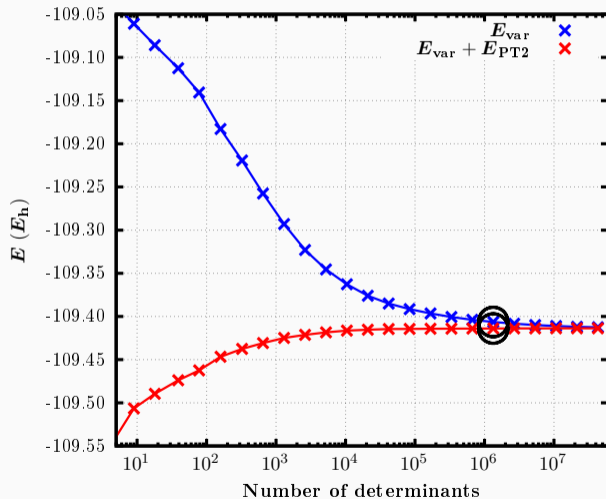
$\text{N}_2$ , aug-cc-pVTZ+3s3p3d





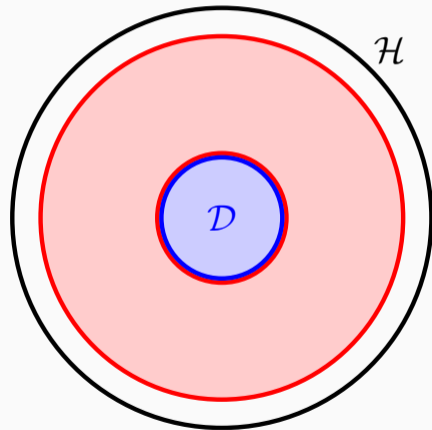
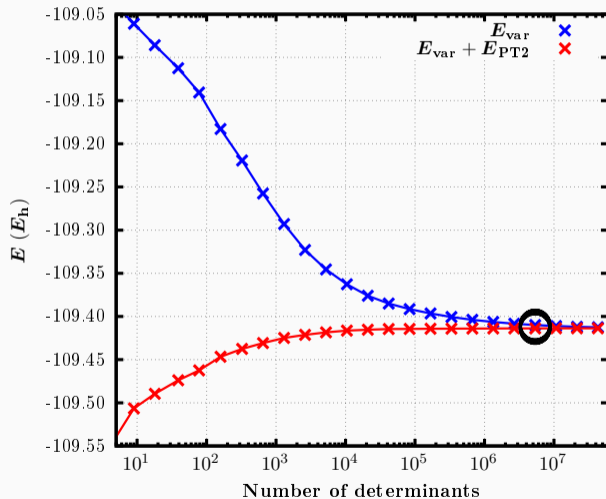
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



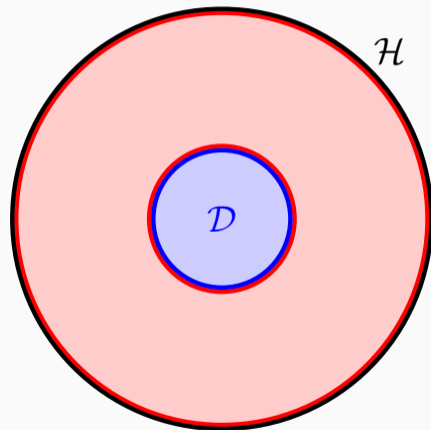
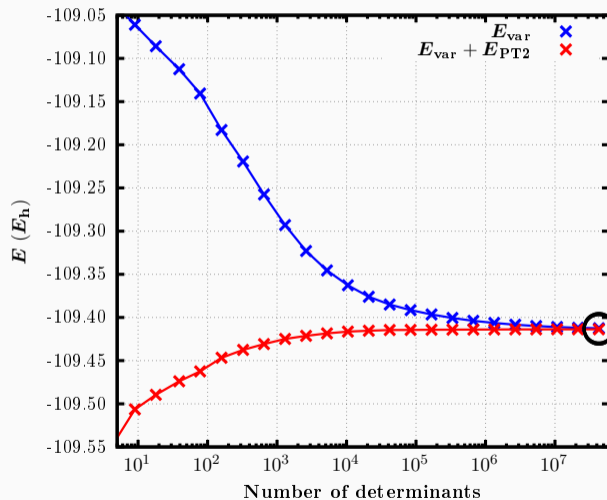
# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



# CIPSI: CI using a Perturbative Selection made Iteratively

$\text{N}_2$ , aug-cc-pVTZ+3s3p3d



## Important

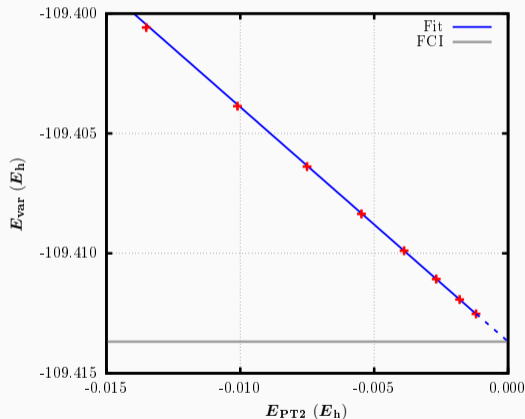
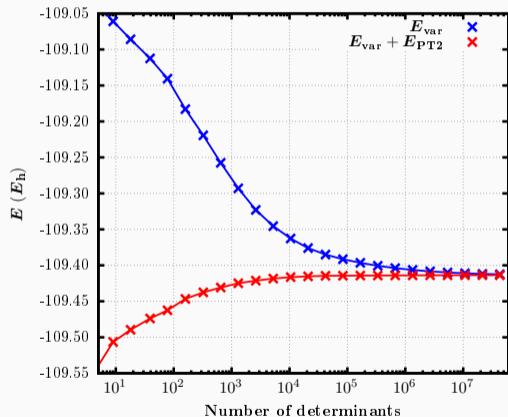
- When all possible the determinants have been included in the space:
  - The FCI solution is obtained

## Important

- When all possible the determinants have been included in the space:
  - The FCI solution is obtained
  - $E_{PT2} = 0$  because of the condition  $|\alpha\rangle \notin \mathcal{D}_n$  in the summation

# Extrapolated FCI energy with CIPSI iterations

exFCI : Extrapolate:  $E = f(E_{PT2})$  at  $E_{PT2} = 0$



# CIPSI: CI using a Perturbative Selection made Iteratively

## Important

- CIPSI is not a *method*, it is a **sparse algorithm** to solve iteratively a CI problem

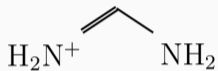
# CIPSI: CI using a Perturbative Selection made Iteratively

## Important

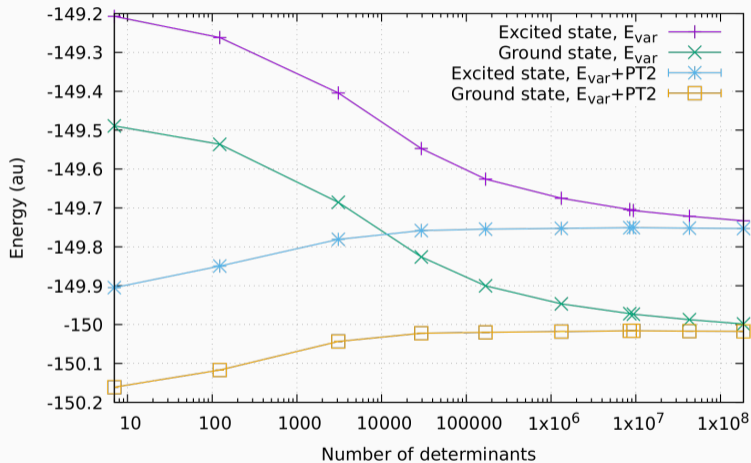
- CIPSI is not a *method*, it is a **sparse algorithm** to solve iteratively a CI problem
  - Every CI problem can be solved with this algorithm
  - Excited states can be obtained



# Excited states

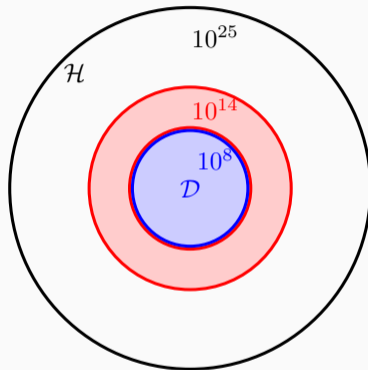
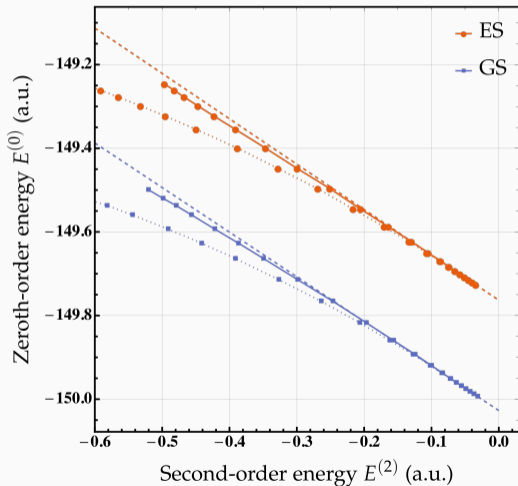


aug-cc-pVDZ  
18 electrons in 111  
orbitals



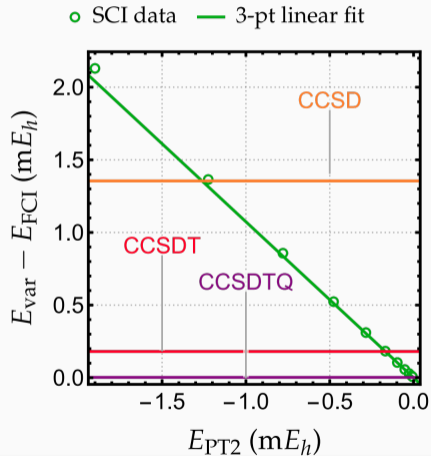
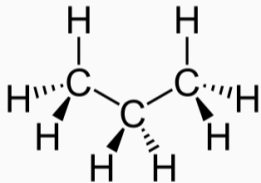
# CIPSI extrapolation

exFCI : Extrapolate  $E = f(E_{\text{PT}2})$  at  $E_{\text{PT}2} = 0$



- $N_{\text{det}}^{\text{FCI}} = 2.5 \times 10^{25}$  (42.4 moles!)
- $10^{14}$  determinants  $\implies$  stochastic algorithm for PT2

# Example: FCI calculation on propane/STO-3G



## Example: FCI calculation on propane/STO-3G

Method	Energy ( $E_h$ )	Error wrt FCI
CCSD	-117.098 767	1.36 m $E_h$
CCSD(T)	-117.099 708	0.41 m $E_h$
CCSDT	-117.099 942 158	0.18 m $E_h$
CCSDTQ	-117.100 120 230	2.45 $\mu E_h$
CIPSI	-117.100 115 87	6.81 $\mu E_h$
CIPSI+PT2	-117.100 122 19	0.49 $\mu E_h$
exFCI	-117.100 122 67(1)	0.01 $\mu E_h$
FCI	-117.100 122 681 461	-

## Example: FCI calculation on propane/STO-3G

### Brute force

- 1.3 trillion determinants  $\implies$  19 TB storage
- 113 hours on 256 nodes (10 240 cores)  $\implies$  1.1 million CPU hours

# Example: FCI calculation on propane/STO-3G

## Brute force

- 1.3 trillion determinants  $\implies$  19 TB storage
- 113 hours on 256 nodes (10 240 cores)  $\implies$  1.1 million CPU hours
- 10 MWh energy:
  - a house for 2.2 years
  - 60 000 km with an electric car



### CIPSI

- Error vs FCI:  $3 \cdot 10^{-6}$  hartree
- 2 million determinants  $\implies$  5 GB storage
- 14 minutes on 1 node (36 cores)  $\implies$  8.4 CPU hours

# Example: FCI calculation on propane/STO-3G

## CIPSI

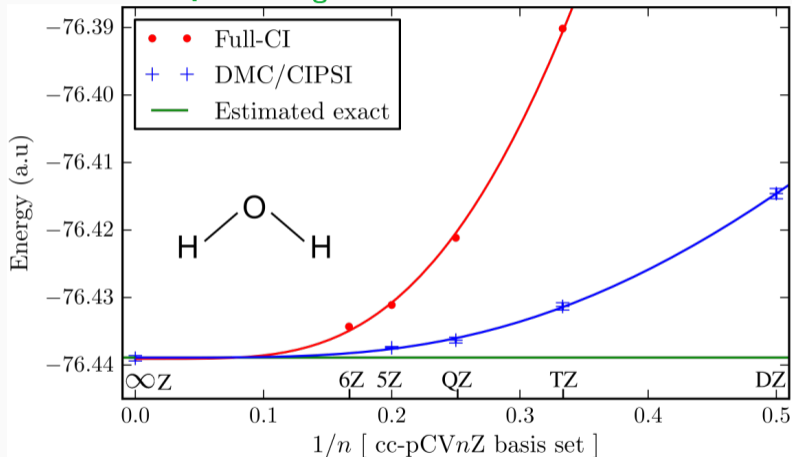
- Error vs FCI:  $3 \cdot 10^{-6}$  hartree
- 2 million determinants  $\implies$  5 GB storage
- 14 minutes on 1 node (36 cores)  $\implies$  8.4 CPU hours
- 74 Wh energy:
  - playing on a PlayStation5 for 20 minutes
  - 500 m with an electric car





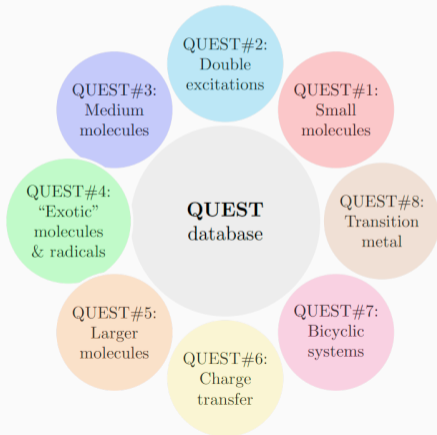
# Some applications

## Post-FCI calculations: QMC using CIPSI trial wave functions



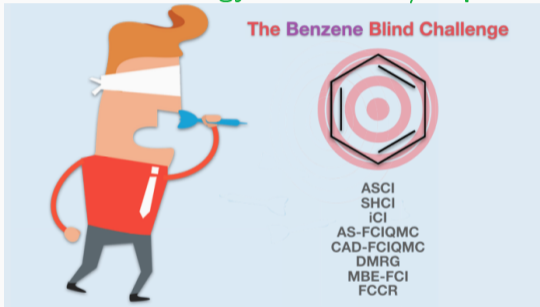
## Quest database: Excited states

*“The QUEST project aims to provide to the community a large set of highly-accurate excitation energies for various types of excited states”*



- ▶ #1: JCTC 14 (2018) 4360
- ▶ #2: JCTC 15 (2019) 1939; arXiv:2403.19597
- ▶ #3: JCTC 16 (2020) 1711
- ▶ #4: JCTC 16 (2020) 3720
- ▶ #5: WIREs 11 (2021) e1517
- ▶ #6: JCTC 17 (2021) 3666
- ▶ #7: JPCA 125 (2021) 10174
- ▶ #8: JCTC 19 (2023) 8782

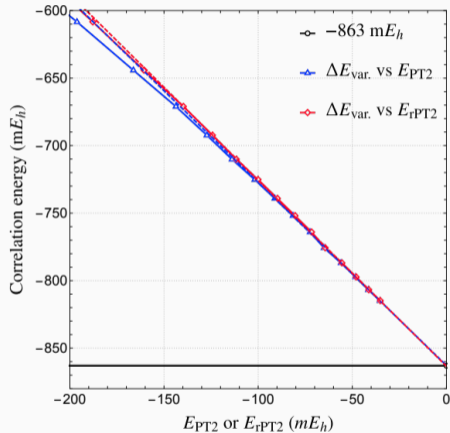
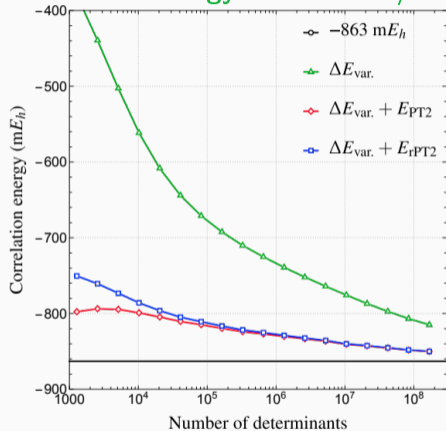
## Correlation energy of Benzene/cc-pVDZ



Eriksen *et al*, doi:10.1021/acs.jpcllett.0c02621

# Some applications

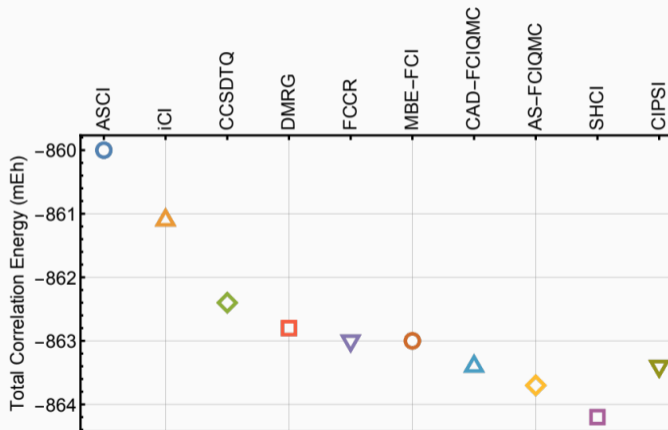
## Correlation energy of Benzene/cc-pVDZ



Loos *et al*, doi:10.1063/5.0027617

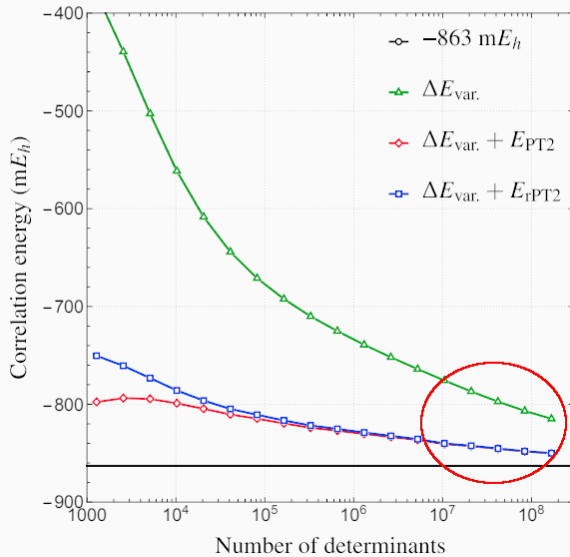
# Some applications

## Correlation energy of Benzene/cc-pVDZ

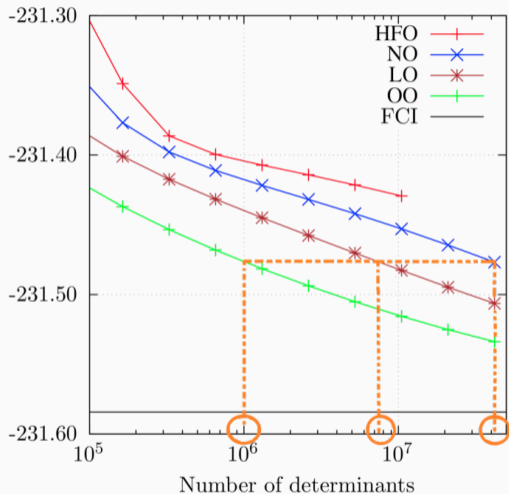


Loos *et al*, doi:10.1063/5.0027617

## Slow convergence of the tail



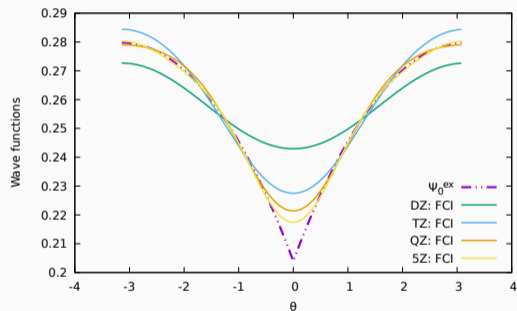
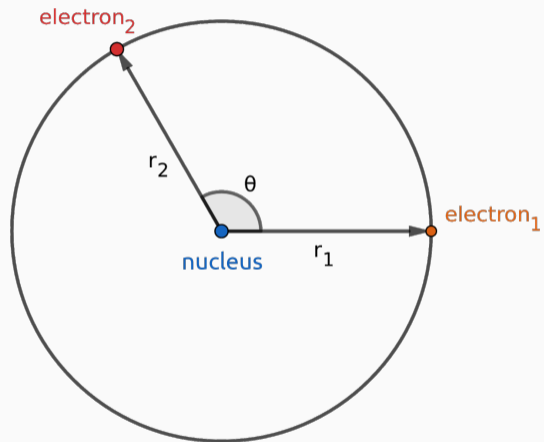
# Orbital optimization



Orbitals	$\times 10^6$ dets	Time
Natural orbitals	44	31 h
Localized orbitals	7	5 h
Optimized orbitals	1	20 min

- Optimized orbitals: Same energy with  $100\times$  less Wh than NO
- Not enough to treat much larger systems

# The Coulomb hole





# Better handling of the Coulomb hole

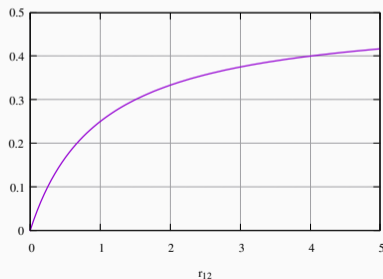
## Quantum Monte Carlo

- Introduce a **Jastrow factor**  $J(r_1, r_2, \dots, r_N)$  in the **wave function**, such that

$$\left. \frac{\partial J}{\partial r_{12}} \right|_{r_{12} \rightarrow 0} = \frac{1}{2}$$

for example:  $J = \frac{1}{2} \sum_{i < j} \frac{r_{ij}}{1 + b r_{ij}},$

$$\Phi = \Psi e^J = \sum_i c_i D_i e^J$$



# Better handling of the Coulomb hole

## Quantum Monte Carlo

- Introduce a **Jastrow factor**  $J(r_1, r_2, \dots, r_N)$  in the **wave function**, such that

$$\left. \frac{\partial J}{\partial r_{12}} \right|_{r_{12} \rightarrow 0} = \frac{1}{2} \quad \text{for example: } J = \frac{1}{2} \sum_{i < j} \frac{r_{ij}}{1 + b r_{ij}}, \quad \Phi = \Psi e^J = \sum_i c_i D_i e^J$$

- Optimize the parameters of the Jastrow factor, and obtain a lower *variational* energy:

$$E = \frac{\langle \Phi | \mathcal{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} = \frac{\langle \Psi e^J | \mathcal{H} | e^J \Psi \rangle}{\langle \Psi e^J | e^J \Psi \rangle}$$

# Better handling of the Coulomb hole

## Quantum Monte Carlo

- Introduce a **Jastrow factor**  $J(r_1, r_2, \dots, r_N)$  in the **wave function**, such that

$$\left. \frac{\partial J}{\partial r_{12}} \right|_{r_{12} \rightarrow 0} = \frac{1}{2} \quad \text{for example: } J = \frac{1}{2} \sum_{i < j} \frac{r_{ij}}{1 + b r_{ij}}, \quad \Phi = \Psi e^J = \sum_i c_i D_i e^J$$

- Optimize the parameters of the Jastrow factor, and obtain a lower *variational* energy:

$$E = \frac{\langle \Phi | \mathcal{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} = \frac{\langle \Psi e^J | \mathcal{H} | e^J \Psi \rangle}{\langle \Psi e^J | e^J \Psi \rangle}$$

- Non-orthogonal basis:  $\langle D_i e^J | e^J D_j \rangle \neq \delta_{ij}$
- Involves  $3N$ -dimensional integrals  $\implies$  Monte Carlo integration

## Transcorrelated (TC) Hamiltonian

- Introduce a **Jastrow factor**  $J(r_1, r_2, \dots, r_N)$  in the **Hamiltonian**, with a similarity transformation:

$$\tilde{\mathcal{H}} \longrightarrow e^{-J} \mathcal{H} e^J$$

- The exact energy is obtained in a complete basis set
- Involves up to three-electron integrals

# Better handling of the Coulomb hole

## Transcorrelated (TC) Hamiltonian

- Introduce a **Jastrow factor**  $J(r_1, r_2, \dots, r_N)$  in the **Hamiltonian**, with a similarity transformation:

$$\tilde{\mathcal{H}} \longrightarrow e^{-J} \mathcal{H} e^J$$

- The exact energy is obtained in a complete basis set
- Involves up to three-electron integrals
- Non-Hermitian operator

- Right- and Left- eigenvectors : 
$$\begin{cases} \tilde{\mathcal{H}}|\Psi_k\rangle &= \tilde{E}_k|\Psi_k\rangle \\ \tilde{\mathcal{H}}^\dagger|\chi_k\rangle &= \tilde{E}_k|\chi_k\rangle \end{cases}$$
- Bi-orthonormal eigenvectors:  $\langle\chi_i|\Psi_j\rangle = \delta_{ij}$ , but  $\langle\Psi_i|\Psi_j\rangle \neq \delta_{ij}$  and  $\langle\chi_i|\chi_j\rangle \neq \delta_{ij}$

# Better handling of the Coulomb hole

## Transcorrelated (TC) Hamiltonian

- Introduce a **Jastrow factor**  $J(r_1, r_2, \dots, r_N)$  in the **Hamiltonian**, with a similarity transformation:

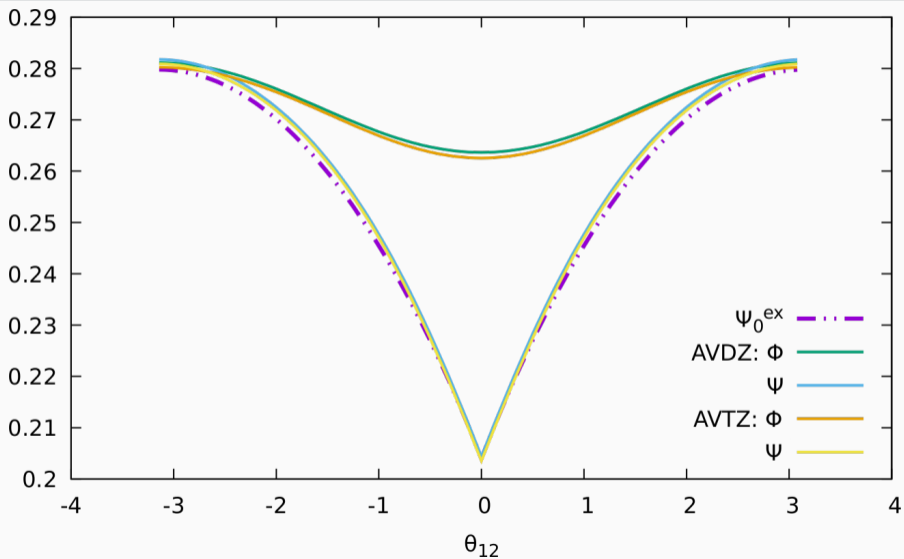
$$\tilde{\mathcal{H}} \longrightarrow e^{-J} \mathcal{H} e^J$$

- The exact energy is obtained in a complete basis set
- Involves up to three-electron integrals
- Non-Hermitian operator
  - *Non-variational* energy:

$$\tilde{E} = \frac{\langle \chi | \tilde{\mathcal{H}} | \Psi \rangle}{\langle \chi | \Psi \rangle} = \frac{\langle \chi | e^{-J} \mathcal{H} e^J | \Psi \rangle}{\langle \chi | \Psi \rangle} = \frac{\langle \chi e^{-J} | \mathcal{H} | e^J \Psi \rangle}{\langle \chi e^{-J} | e^J \Psi \rangle}$$

- Orthogonal determinant basis:  $\langle D_i e^{-J} | e^J D_j \rangle = \langle D_i | D_j \rangle = \delta_{ij}$

## Better handling of the Coulomb hole



## Implementation in Quantum Package

- 1,2,3-e Integrals involving the Jastrow factor
- Modified Slater-Condon rules
- Modified PT2 contributions using

$$\frac{\langle \chi_n | \tilde{\mathcal{H}} | \alpha_r \rangle \langle \alpha_l | \tilde{\mathcal{H}} | \Psi_n \rangle}{\tilde{E}(\Psi_n) - \langle \alpha_l | \tilde{\mathcal{H}} | \alpha_r \rangle}$$

- Non-Hermitian Davidson diagonalization
- TC-SCF: left and right bi-orthonormal molecular orbitals  $\implies$  different left and right determinant bases

Ammar *et al*, doi:10.1063/5.0115524, doi:10.1063/5.0163831 and doi:10.1021/acs.jctc.3c00257



# Improved energies

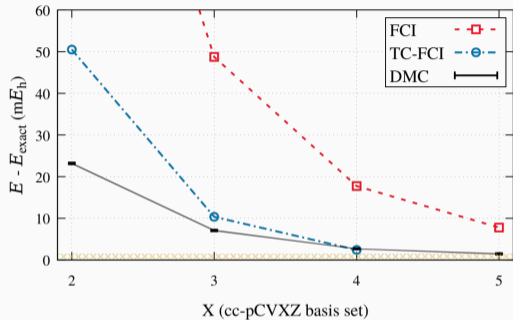


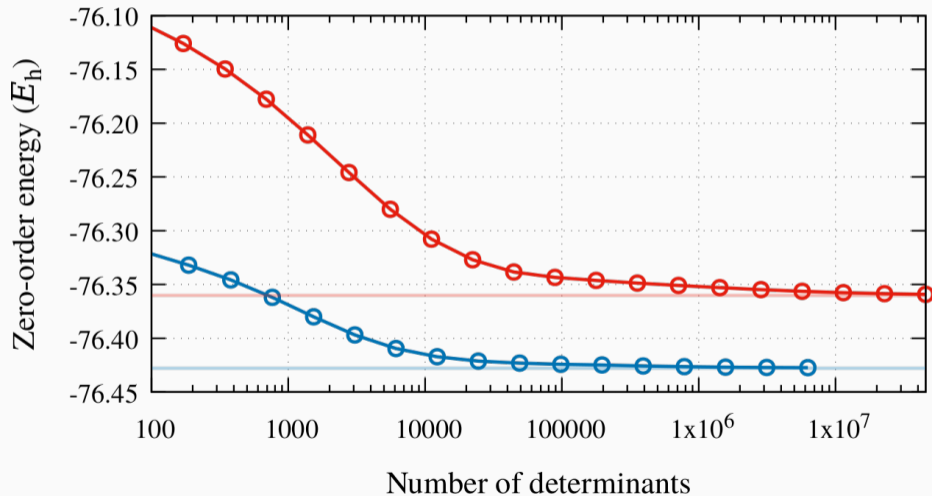
FIG. 9. Error (in  $mE_h$ ) in the total energy of  $\text{H}_2\text{O}$  computed at the FCI, TC-FCI, and FN-DMC levels with the cc-pCVXZ family of basis sets with respect to the estimated exact energy ( $-76.4389 E_h$ ). The shaded yellow region corresponds to  $1.5 mE_h$  accuracy.

TABLE III. Total energies (in  $E_h$ ) computed at the FCI, TC-FCI, and FN-DMC levels of theory for the water molecule in the cc-pCVXZ family of basis sets. The estimated exact energy is  $-76.4389 E_h$ .<sup>149</sup>

	FCI <sup>a</sup>	TC-FCI <sup>b</sup>	FN-DMC <sup>c</sup>
cc-pCVDZ	-76.282 9	-76.388 4	-76.415 71(20)
cc-pCVTZ	-76.390 2	-76.428 6	-76.431 82(19)
cc-pCVQZ	-76.421 2	-76.436 5	-76.436 22(14)
cc-pCV5Z	-76.431 1		-76.437 44(18)

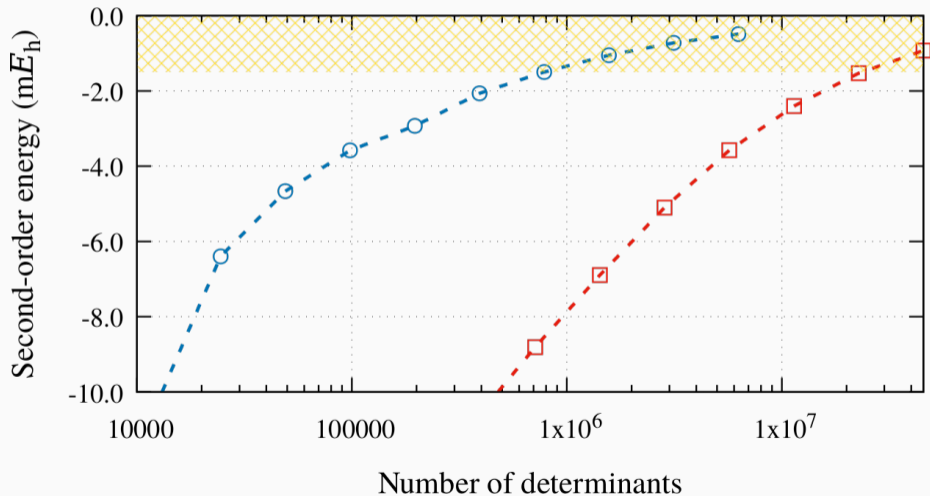
# Acceleration of the convergence

H<sub>2</sub>O@cc-pVQZ



# Acceleration of the convergence

H<sub>2</sub>O@cc-pVQZ



# Ionization potentials

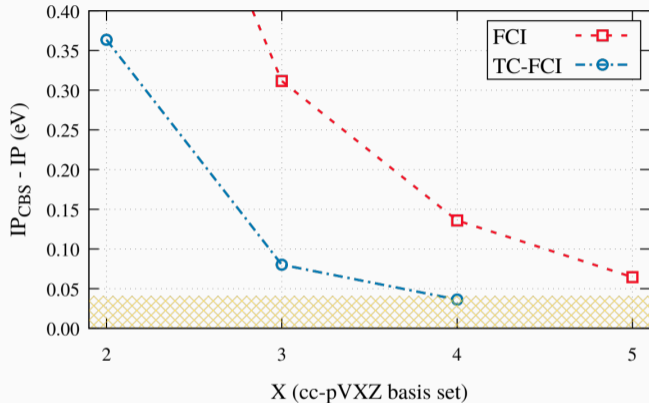


FIG. 10. Error (in eV) in the IP of H<sub>2</sub>O computed at the FCI and TC-FCI levels with the cc-pVXZ family of basis sets with respect to the CBS estimate (12.80 eV). The shaded yellow region corresponds to  $1.5 mE_h$  accuracy.

## CIPSI

- Less energy-consuming algorithm than brute-force FCI
- The Hilbert space grows exponentially fast with tiny contributions
- Current limit for chemical accuracy is close to benzene/cc-pVDZ

## CIPSI

- Less energy-consuming algorithm than brute-force FCI
- The Hilbert space grows exponentially fast with tiny contributions
- Current limit for chemical accuracy is close to benzene/cc-pVDZ

## Beyond CIPSI

- Necessity to:
  - accelerate the convergence of CIPSI

## CIPSI

- Less energy-consuming algorithm than brute-force FCI
- The Hilbert space grows exponentially fast with tiny contributions
- Current limit for chemical accuracy is close to benzene/cc-pVDZ

## Beyond CIPSI

- Necessity to:
  - accelerate the convergence of CIPSI
  - reduce the size of the Hilbert space

## CIPSI

- Less energy-consuming algorithm than brute-force FCI
- The Hilbert space grows exponentially fast with tiny contributions
- Current limit for chemical accuracy is close to benzene/cc-pVDZ

## Beyond CIPSI

- Necessity to:
  - accelerate the convergence of CIPSI
  - reduce the size of the Hilbert space
- Using the TC Hamiltonian shows promising results



# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems

# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large

# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large
  - Is quantum computing really going to be a game-changer for quantum chemistry?

# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large
  - Is quantum computing really going to be a game-changer for quantum chemistry?

## Energy minimization

- 500 000 CPU hours  $\equiv$  electricity needed by a house for 1 year

# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large
  - Is quantum computing really going to be a game-changer for quantum chemistry?

## Energy minimization

- 500 000 CPU hours  $\equiv$  electricity needed by a house for 1 year
  - So many points on curves?

# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large
  - Is quantum computing really going to be a game-changer for quantum chemistry?

## Energy minimization

- 500 000 CPU hours  $\equiv$  electricity needed by a house for 1 year
  - So many points on curves?
  - Do we need all those digits?

# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large
  - Is quantum computing really going to be a game-changer for quantum chemistry?

## Energy minimization

- 500 000 CPU hours  $\equiv$  electricity needed by a house for 1 year
  - So many points on curves?
  - Do we need all those digits?  $\implies$  Sparse approaches (SCI, DLPNO-CCSD, Stochastic methods)

# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large
  - Is quantum computing really going to be a game-changer for quantum chemistry?

## Energy minimization

- 500 000 CPU hours  $\equiv$  electricity needed by a house for 1 year
  - So many points on curves?
  - Do we need all those digits?  $\implies$  Sparse approaches (SCI, DLPNO-CCSD, Stochastic methods)
  - Do we really need double precision?



# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large
  - Is quantum computing really going to be a game-changer for quantum chemistry?

## Energy minimization

- 500 000 CPU hours  $\equiv$  electricity needed by a house for 1 year
  - So many points on curves?
  - Do we need all those digits?  $\implies$  Sparse approaches (SCI, DLPNO-CCSD, Stochastic methods)
  - Do we really need double precision?
  - Do we really need to do always larger simulations?

# Take-home messages and open questions

## Full CI

- FCI energies can be routinely obtained for small systems
  - FCI not as precise as we think: basis set errors are large
  - Is quantum computing really going to be a game-changer for quantum chemistry?

## Energy minimization

- 500 000 CPU hours  $\equiv$  electricity needed by a house for 1 year
  - So many points on curves?
  - Do we need all those digits?  $\implies$  Sparse approaches (SCI, DLPNO-CCSD, Stochastic methods)
  - Do we really need double precision?
  - Do we really need to do always larger simulations?
    - Propane: removing 2 hydrogens  $\implies$  15 $\times$  less determinants